STAT 205A - Problem Set 06

Walid Krichene (23265217)

October 15, 2013

(6.1) Let (X_i) be independent, $S_n = \sum_{i=1}^n X_i$, $S_n^* = \max_{i\leq n} |S_i|$. Prove that

$$P(S_n^* > 2a) \leq \frac{P(|S_n| > a)}{\min_{j\leq n} P(|S_n - S_k| \leq a)}, a > 0$$

proof First, let us partition the event

$$\{S_n^* > 2a\} = \cup_{j=1}^n A_k$$

where $A_k = \{S_k > 2a \text{ and } S_i \leq 2a \ \forall i < j\}$. Here the A_k's are disjoint.

Now we observe (given the hint) that for all j,

$$A_k(\omega) \text{ and } |S_n(\omega) - S_k(\omega)| \leq a \Rightarrow |S_n(\omega)| > a$$

this follows from the triangle inequality since $|S_k| \leq |S_n - S_k| + |S_n|$

$$|S_n(\omega)| \geq |S_k(\omega)| - |S_n(\omega) - S_k(\omega)| > 2a - a = a$$

therefore for all j,

$$P(A_k \text{ and } |S_n - S_k| \leq a) \leq P(|S_n| > a)$$

here A_k and $S_n - S_k$ are disjoint, since A_k is a function of $\{X_i\}_{i\leq j}$ and $S_n - S_k$ is a function of $\{X_i\}_{i>j}$. Therefore we have

$$P(A_k) \min_{j\leq n} P(|S_n - S_j| \leq a) \leq P(A_k)P(|S_n - S_k| \leq a) = P(A_k \text{ and } |S_n - S_k| \leq a) \leq P(|S_n| > a)$$

Summing over k, we have

$$\sum_{k=1}^n P(A_k) \min_{j\leq n} P(|S_n - S_j| \leq a) \leq P(|S_n| > a)$$

i.e.

$$P(S_n^* > 2a) \min_{j\leq n} P(|S_n - S_j| \leq a) \leq P(|S_n| > a)$$

(6.2) In the setting of the previous question, prove

(i) if $\lim_{n\to\infty} S_n$ exists in probability, then the limit exists a.s.

proof $\lim_n S_n$ exists in probability means that there exists a function $l(\omega)$ such that for all $\epsilon > 0$, $P(|S_n - l| > \epsilon)$ converges to 0 as $n \to \infty$. This implies in particular that S_n is Cauchy in probability, i.e. for all $\epsilon > 0$,

$$\lim_{N\to\infty} \sup_{n,m\geq N} P(|S_n - S_m| > \epsilon) = 0 \tag{1}$$

this follows from the fact that $|S_n - S_m| \leq |S_n - l| + |S_m - l|$, thus

$$|S_n - l| \leq \epsilon/2 \text{ and } |S_m - l| \leq \epsilon/2 \Rightarrow |S_n - S_l| \leq \epsilon$$

or, taking complements,

$$|S_n - S_m| > \epsilon \Rightarrow |S_n - l| > \epsilon/2 \text{ or } |S_m - l| > \epsilon/2$$

therefore

$$P(|S_n - S_m| > \epsilon) \leq P(|S_n - l| > \epsilon/2 \text{ or } |S_m - l| > \epsilon/2) \leq P(|S_n - l| > \epsilon/2) + P(|S_m - l| > \epsilon/2)$$

and both terms converge to 0 as m, n tend to infinity.

Now fix $\epsilon > 0$ and consider the nested sequence of events

$$A_N = \cup_{m,n\geq N}\{|S_m - S_n| > \epsilon\}$$

Then we have

$$\begin{aligned}
P(|S_m - S_n| \text{ infinitely often }) &= P(\cap_N A_N) \\
&= \lim_{N\to\infty} P(A_N) && \text{since the sequence is nested} \\
&\leq \lim_{N\to\infty} \sup_{m,n\geq N} P(|S_m - S_n| > \epsilon) \\
&= 0 && \text{by (1)}
\end{aligned}$$

Therefore

$$P(|S_m - S_n| > \epsilon \text{ infinitely often}) = 0$$

This is true for all $\epsilon > 0$, and in particular for all ϵ positive rational. Since $\mathbb{Q}$ is countable, we have

$$P(\exists \epsilon \in \mathbb{Q}_+ \text{ such that } |S_m - S_n| > \epsilon \text{ infinitely often}) = 0$$

i.e.

$$P(\forall \epsilon \in \mathbb{Q}_+, |S_m - S_n| \leq \epsilon \text{ ultimately}) = 1$$

therefore S_n is Cauchy a.s., and it follows that S_n converges a.s. (for any ω such that $S_n(\omega)$ is Cauchy, $S_n(\omega)$ converges)

(ii) if the (X_i) are identically distributed and if $\frac{S_n}{n} \to 0$ in probability, then $\frac{\max_{m\leq n} S_m}{n} \to 0$ in probability.

proof Let $\epsilon > 0$. We seek to show that $P(\frac{1}{n}|\max_{m\leq n} S_m| > \epsilon) \to 0$. Since $|\max_{m\leq n} S_m| \leq \max_{m\leq n} |S_m| = S_n^*$, it suffices to show that $P(\frac{1}{n}S_n^* > \epsilon) \to 0$. From problem 1, we have

$$P(\frac{1}{n}S_n^* > \epsilon) \leq \frac{P(\frac{1}{n}|S_n| > \epsilon/2)}{\min_{j\leq n} P(\frac{1}{n}|S_n - S_j| \leq \epsilon/2}$$

where the numerator, $P(\frac{1}{n}|S_n| > \epsilon/2)$ converges to 0 by assumption on S_n. Therefore it suffices to lower bound the denominator $\min_{j\leq n} P(\frac{1}{n}|S_n - S_j| \leq \epsilon/2) \leq M$ for some constant M (uniformly in n).

We have by the triangle inequality $|S_n - S_j| \leq |S_n| + S_j|$ thus $|S_n - S_j| > \epsilon/2 \Rightarrow |S_n| > \epsilon/4$ or $|S_j| > \epsilon/4$, thus for all n and all $j \leq n$

$$\begin{aligned} P(\frac{1}{n}|S_n - S_j| \leq \epsilon/2) &= 1 - P(\frac{1}{n}|S_n - S_j| > \epsilon/2) \\ &\geq 1 - P(\frac{1}{n}|S_n| > \epsilon/4 \text{ or } \frac{1}{n}|S_j| > \epsilon/4) \\ &\geq 1 - P(\frac{1}{n}|S_n| > \epsilon/4) - P(\frac{1}{n}|S_j| > \epsilon/4) \end{aligned}$$

Now fix $0 < \delta < 1/2$. Since $\frac{|S_n|}{n}$ converges to 0 in probability, there exists N such that for all $n \geq N$, $P(\frac{1}{n}|S_n| > \epsilon/4) \leq \delta$. Now for the last term, we have

- for all $j \geq N$
$$P(\frac{1}{n}|S_j| > \epsilon/4) \leq P(\frac{1}{j}|S_j| > \epsilon/4) \leq \delta$$
- for all $j < N$
$$P(\frac{1}{n}|S_j| > \epsilon/4) \to 0 \text{ as } n \to \infty$$

  Therefore there exists N' such that for $n > N'$, $P(\frac{1}{n}|S_j| > \epsilon/4) \leq \delta$ for all $j < N$.

Combining the above inequalities, we have for all $n \geq \max(N, N')$, and for all $j \leq n$

$$\begin{aligned} P(\frac{1}{n}|S_n - S_j| \leq \epsilon/2) &\geq 1 - P(\frac{1}{n}|S_n| > \epsilon/4) - P(\frac{1}{n}|S_j| > \epsilon/4) \\ &\geq 1 - 2\delta > 0 \end{aligned}$$

And using the result of problem (1)

$$P(\frac{1}{n}S_n^* > \epsilon) \leq \frac{1}{1-2\delta} P(\frac{1}{n}|S_n| > \epsilon/2)$$

and the righthand side converges to 0 as $n \to \infty$, which proves the result

(6.3) Let (X_i) be i.i.d., taking values in $\{-1, 1, 3, 7, 15, \dots\}$, such that

$$P(X = 2^k - 1) = \frac{1}{k(k+1)2^k}, k \geq 1$$

(this implicitly specifies $P(X = -1)$).

(i) Show that $\mathrm{E}\, X_1 = 0$

proof We have

$$P(X = -1) = 1 - \sum_{k \geq 1} \frac{1}{k(k+1)2^k} \tag{2}$$

The expectation is

$$\begin{aligned}
\mathrm{E}\, X &= \sum_{k \geq 0} (2^k - 1) P(X = 2^k - 1) \\
&= -1.P(X = -1) + \sum_{k \geq 1} (2^k - 1) \frac{1}{k(k+1)2^k} \\
&= -P(X = -1) + \sum_{k \geq 1} \frac{2^k}{k(k+1)2^k} + P(X = -1) - 1 \qquad \text{using (2)} \\
&= \sum_{k \geq 1} \frac{1}{k(k+1)} - 1 \\
&= 0
\end{aligned}$$

using the fact that $\sum_{k \geq 1} \frac{1}{k(k+1)} = \sum_{k \geq 1} \frac{1}{k} - \frac{1}{k+1} = 1$ (telescoping partial sums).

(ii) Show that for all $\alpha < 1$,[1]

$$P\left(S_n < -\frac{\alpha n}{\log_2 n}\right) \to 1$$

proof We use the result from Durrett, called Weak Law for Triangular Arrays (Theorem 2.2.6), stated below (adapted to the notation of this problem)

Theorem 1 *Let (X_i) be i.i.d. Let $b_n > 0$ with $b_n \uparrow \infty$. Suppose that*

(i) $\lim_{n \to \infty} \sum_{i=1}^n P(|X_i| > b_n) = 0$

(ii) $\lim_{n \to \infty} \frac{1}{b_n^2} \sum_{i=1}^n \mathrm{E}[X_i^2 1_{X_i \leq b_n}] = 0$

Then

$$\frac{S_n - a_n}{b_n} \xrightarrow{P} 0$$

where $a_n = \sum_{i=1}^n \mathrm{E}[X_i 1_{X_i \leq b_n}]$

Let $m(n) = \min\{m : \frac{1}{2^m m^{\frac{3}{2}}} \leq \frac{1}{n}\}$ and define $b_n = 2^{m(n)}$. We have $m(n) \uparrow \infty$, thus $b_n \uparrow \infty$ as $n \to \infty$. We first show conditions (i) and (ii):

[1]This is sometimes described as "an unfair, fair game". It shows that the conclusion of the SLLN and the "recurrence of sums" theorem can't be strengthened much

(i) We have

$$\begin{aligned}
\sum_{i=1}^{n} P(X_i > b_n) &= nP(X > b_n) \\
&= n \sum_{k=m(n)+1}^{\infty} \frac{1}{k(k+1)2^k} \\
&\leq \frac{n}{m(n)^2} \sum_{k=m(n)+1}^{\infty} \frac{1}{2^k} && \text{using the fact that } k \geq m(n) \\
&= \frac{n}{m(n)^2} \frac{1}{2^{m(n)}} \\
&\leq \frac{1}{m(n)^{\frac{1}{2}}} && \text{using the fact that } \frac{n}{2^{m(n)}} \leq m(n)^{\frac{3}{2}}
\end{aligned}$$

which converges to 0 as $n \to \infty$ (since $\lim_n m(n) = \infty$).

(ii) We have

$$\begin{aligned}
\frac{1}{b_n^2} \sum_{i=1}^{n} \mathrm{E}[X_i^2 1_{X_i \leq b_n}] &= \frac{n}{b_n^2} \mathrm{E}[X^2 1_{X \leq b_n}] \\
&= \frac{n}{b_n^2} \left((-1)^2 P(X=-1) + \sum_{k=1}^{m(n)} P(X = 2^k - 1)(2^k - 1)^2 \right) \\
&= \frac{n}{b_n^2} \left(1 - \sum_{k=1}^{\infty} \frac{1}{k(k+1)2^k} + \sum_{k=1}^{m(n)} \frac{2^{2k} + 1 - 2^{k+1}}{k(k+1)2^k} \right) \\
&= \frac{n}{b_n^2} \left(1 - \sum_{k=1}^{\infty} \frac{1}{k(k+1)2^k} + \sum_{k=1}^{m(n)} \frac{2^k}{k(k+1)} + \sum_{k=1}^{m(n)} \frac{1}{k(k+1)2^k} - \sum_{k=1}^{m(n)} \frac{2}{k(k+1)} \right) \\
&= \frac{n}{b_n^2} \left(1 - \sum_{k=m(n)+1}^{\infty} \frac{1}{k(k+1)2^k} + \sum_{k=1}^{m(n)} \frac{2^k}{k(k+1)} - 2 \sum_{k=1}^{m(n)} (\frac{1}{k} - \frac{1}{k+1}) \right) \\
&= \frac{n}{b_n^2} \left(1 - \sum_{k=m(n)+1}^{\infty} \frac{1}{k(k+1)2^k} + \sum_{k=1}^{m(n)} \frac{2^k}{k(k+1)} - 2(1 - \frac{1}{m(n)+1}) \right) \\
&\leq \frac{n}{b_n^2} \left(\sum_{k=1}^{m(n)} \frac{2^k}{k(k+1)} + \frac{2}{m(n)} \right)
\end{aligned}$$

the second term converges to 0 by definition of $m(n)$ since $\frac{n}{b_n^2 m(n)} \leq \frac{m(n)^{\frac{1}{2}}}{2^{m(n)}}$ which converges to 0. To show the first terms converges to 0, we can write

$$\begin{aligned}
\frac{n}{b_n^2} \sum_{k=1}^{m(n)} \frac{2^k}{k(k+1)} &= \frac{n}{b_n^2} \left(\sum_{k=1}^{m(n) - 2\log_2 m(n)} \frac{2^k}{k(k+1)} + \sum_{k=1}^{m(n) - 2\log_2 m(n)} \frac{2^k}{k(k+1)} \right) \\
&\leq \frac{n}{b_n^2} \left(\frac{2^{m(n)}}{m(n)^2} + \frac{2^{m(n)}}{m(n)^2} 2\log_2 m(n) + 1 \right)
\end{aligned}$$

and each term converges to 0 by definition of $m(n)$.

Therefore by the Theorem above, $\frac{S_n - a_n}{b_n} \to 0$ in probability. Now it suffices to show that $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$ and $b_n \sim \frac{\log_2 n}{n}$

We have

$$\begin{aligned}
a_n &= \sum_{i=1}^{n} \mathrm{E}[X_i 1_{X_i \le b_n}] \\
&= n\,\mathrm{E}[X 1_{X \le b_n}] \\
&= n \left(\sum_{k=1}^{m(n)} \frac{2^k - 1}{k(k+1)2^k} - (1 - \sum_{k=1}^{\infty} \frac{1}{k(k+1)2^k}) \right)
\end{aligned}$$